US ARMY
LABORATORY COMMAND

CONTRACT REPORT NUMBER 36-89
PREPARED FOR THE

HUMAN ENGINEERING LABORATORY

BY

University of Massachusetts, Amherst, Massachusetts

PROBING THE MENTAL MODELS OF SYSTEM STATE CATEGORIES WITH MULTIDIMENSIONAL SCALING

Dec 1987

Bruce G. Coury
Monica C. Zubritzky
V. Grayson Cuqlock-Knopp

Contract DAAA15-85-K-0010

US ARMY HUMAN ENGINEERING LABORATORY
ABERDEEN PROVING GROUND, MARYLAND
HEL

ABERDEEN PROVING GROUND, MARYLAND 21005-5001

Probing the Mental Models of System State Categories with Multidimensional Scaling

by

Bruce G. Coury, Monica C. Zubritzky
and V. Grayson Cuqlock-Knopp

December 1987

FOOTNOTES

Manuscript received . . . This research was supported by U.S. Army Human Engineering Laboratory Contract No. DAAA15-85-K-0010 to the first author.

B.G. Coury and M.C. Zubritzky are with the Department of Industrial Engineering and Operations Research, University of Massachusets, Amherst, MA 01003, USA.

V.G. Cuqlock-Knopp is with the Behavioral Research Division of the U.S. Army Human Engineering Laboratory, Aberdeen Proving Grounds, MD, 21005, USA.

ABSTRACT

Identifying the underlying decision criteria used by operators to classify system state and revealing the way in which that information is internally represented by individual operators is one of the major challenges facing designers of decision aids for process plants. This research describes the use of multidimensional scaling (MDS) to probe the structure and composition of the mental models of used by operators to identify system state, and evaluate the impact of different display representations on those models. Twenty subjects were trained to classify instances of system data. Pairwise similarity ratings of instances of system data were analyzed by MDS to reveal the dominant dimensions used by operators. Results showed that significant individual differences emerged, and that the dimensions used by subjects were also a function of the type of display representation.

INTRODUCTION

Identification and recognition of system state is fundamental to the effective supervision of a complex system. In supervisory tasks where control decisions are based on multidimensional data and information, the operator's ability to map the values of critical system variables to known definitions of system state is the prerequisite step to selecting the best course of action [1].

Accurate identification of system state presupposes, however, that the operator possess a well developed internal model of the criteria defining system state categories [1], [2], [3], [4]. Unfortunately, designers of user interfaces and decision aids for complex systems are hindered by an incomplete understanding of both the knowledge used by operators to assess the status of a system, and the impact of the operator's mental model on performance. That incomplete understanding is due, in part, to the paucity of techniques for assessing the operator's internal model and relating the composition of that model to performance. The primary purpose of the research presented in this paper is to explore the use of multidimensional scaling (MDS) as a method for probing the mental models of system state categories.

Fundamental to this research is the notion that identification of system state is a categorization process. In our approach to system state identification, a system state category, S_i, is defined by a set of system state attributes, $\{a_1, \ldots, a_n\}$, with prespecified ranges of values for a set of attributes delimiting membership in a specific state category. Therefore, a system state category is defined by:

$$S_i = \{w_1 a_1, \ldots, w_n a_n\}$$

where the weights, w_j, for each attribute determine the importance of that attribute to specifying a particular system state. In this research, process variables serve as state attributes. Thus, an instance of system state is a specific set of values for the process variables from a particular state category. Thus, classification of an instance of system data requires that the operator map data from each of the relevant process variables to a set of decision criteria defining a system state category [5], [6], [7].

Knowledge of the mapping of system data to system state categories is contained within the operator's mental model of the decision task. The model must contain, at a minimum, knowledge of the attributes which define system state, the relative importance of those attributes in identifying a specific state, and the decision rules used by the operator to map attributes to state categories. The role of a mental model is to organize that knowledge in such a way as to facilitate classification of system state. The operator's performance, then, is dependent upon his or her ability to match the data from each information source to an internal model defining category membership, and accurately classify the state of the system. A well developed mental model, then, is composed of the knowledge necessary for making an accurate and timely assessment of the status of a system.

Using Multidimensional Scaling

MDS is used in this research to probe the structure and composition of the internal models used by operators to identify system state. MDS is method based on least squares regression analysis which relates the judged dissimilarities between objects to derived distances in a multidimensional geometric space [8], [9], [10], [11]. According to the MDS approach there should be a reliable relationship between the cognitive construct "similarity" and distance in a geometric space. In general, objects which possess common features or attributes should be rated as similar and be located in close proximity in space (in the context of this research, the objects are instances of system data). Conversely, dissimilar objects should be distant from each other in that same space. The technique has been used in a number of contexts, including concept learning [12], judgmental process in analogical reasoning [13], the evolution of conceptual structure [14], and visual inspection [15].

The MDS algorithm used in this study [10], [11], represents the similarities between instances of system data as euclidean distances in a n-dimensional space. The method generates spatial configurations of a set of rated objects, thus providing a number of potentially useful measures of distance and structure for probing conceptual models. In addition to the spatial configurations, the algorithm provides the coordinate position of each object in the n-dimensional space, and a measure of goodness-of-fit called "Stress" to evaluate the configuration.

Stress is defined as the amount of variation between the theoretical distances represented by the similarity rating data and the distances calculated by the algorithm. In this experiment, solutions were generated for one to five dimensions. The objective was to find the number of dimensions which minimized the value of Stress; thus, the best fit is defined as that number of dimensions where Stress reaches an asymptotic level. The spatial configurations representing the best fit can then be visually examined to determine the extent of clustering and interpret the meaning of the dimensions. Thus, the spatial configurations reveal the attributes used by operators to assess the similarity between instances of system data.

The coordinate points for each object can be used to derive measures of structure and distance; in this research, structural ratios and directed distances were the measures of interest. The structural ratio is a measure of the degree of categorical clustering of a set of objects. The ratio is based on the presupposition that items which are clustered together in the same geometric space are in some way categorically similar. The ratio relates the mean intracategory distances to the mean intercategory distances [14], [15]. As category structure increases, the distance between members of the same category decreases, the distance between categories increases, and the structural ratio approaches zero. Thus, the ratio can be used to characterize the strength of a dimension for a particular operator. Dimensions which produce the most strength (i.e., a structural ratio closest to zero) are the most dominant in the operator's mental model and are the most likely to be reflected in his/her classification performance. Thus, the assessment of a dimension's dominance within a mental model is critical to finding a link between a mental model and performance, and the coordinate distances provide a means for weighting the importance of each dimension.

Unfortunately, the structural ratio is sensitive only to dimensions which produce clustering, and cannot characterize directed distances between instances in a multidimensional space. As a result, measures of directed distances must also be derived by calculating the euclidean distance between instances across all dimensions of interest, using an object or set of objects as a referent for the vector distance.

The Effect of Display Type

An important consideration in the development of an operator's mental model of system state categories is the manner in which system data is presented during training. There has been a number of recent studies concerned with the issue of the physical (display) representation of system data [16], [17], [18], [19], [20]. The results of those studies indicate quite clearly that the choice of a display is dependent upon the underlying statistical properties of the task, the type of task (e.g., fault detection or fault diagnosis), and the degree of uncertainty involved in identifying the state of the system. None of those studies have directly assessed the impact of display representation on the operator's internal model of the task, or considered the real possibility that the operator's knowledge of a system will be fundamentally influenced by the way in which data is displayed This research specifically addresses that issue.

Research Objectives

This research has four very specific objectives related to the use of multidimensional scaling (MDS) and the design of user interfaces. Those objectives are: 1) explore the use of MDS as a method for revealing the composition and structure of an operator's mental of state categories; 2) determine the relationship between an operator's internal model and classification performance; 3) evaluate the differences between the models revealed by the group MDS analysis and the models used by individuals; and 4) determine the extent to which display representations affect the composition of the operator's mental model.

Objectives 1 through 3 are primarily concerned with evaluating the feasibility of MDS as a tool for assessing operator knowledge of state categories. Since the construction of effective decision aiding and support systems rely upon accurate a priori knowledge of relevant decision variables, there is a need for developing knowledge acquisition tools which can rapidly identify those variables and simultaneously evaluate individual operators.

Objective 4 addresses an important interface design issue for control systems; given that system data must be displayed on a screen, how does a specific type of display affect the operator's knowledge of system state categories? In this research, MDS will be used to compare the dominant dimensions of two groups of operators using different types of displays. The outcome of the MDS analysis will reveal the differences, if any, in the set of attributes used by the operators viewing the same data displayed in physically different forms.

The objectives of the research required that a simulated task be used which captured the essence of the cognitive demands placed on operators when identifying system state. The multidimensional decision task used in previous research [18], [19], [20], was used in this study. The task required the operator to integrate information from a number of sources (in this case, process variables), and use that information to identify (classify) the state of the system. Thus, the operator had to learn, during training, the mapping of values of process variables to system state categories.

To probe the mental models of system state categories, operators rated the similarity of a subset of instances of system data after learning. The classification training task allowed the mental model to develop, and the similarity ratings provided the data necessary for characterizing the composition and structure of the operator's internal model. Since the focus of this study was on well-developed mental models, data is reported on only those participants in the study who were able to accurately identify system state.

METHODS and PROCEDURES

Subjects

Twenty students from the University of Massachusetts served as subjects in this experiment. Subjects were paid $5/hour for their participation. The 20 subjects represent only those who reached a prespecified criterion after training.

Apparatus

A DEC Pro 300 series microcomputer was used for presenting instances of system state and recording subject performance data. The stimuli were presented on a high resolution graphics display. A second monitor was used to present feedback to subject responses.

System State Categories

System state categories were defined in this experiment as ranges of values along four dimensions. Each dimension represented one of four process variables; a specific range of values for each process variable was combined to define a state category. The ranges of values of the variables for each of the four system state categories is shown in Table 1.

INSERT TABLE 1

ABOUT HERE

The ranges of values for each state category define the correlational structure [5]. For instance, State Categories 1 and 3 share the same range of values for the third and fourth process variables, as do State Categories 2 and 4. For the classification task, then, those values provide the information necessary to narrow the possibilities to a pair of system states. To distinguish between the chosen pair, the information provided by the first and second process variables is needed. The correlational structure of this task makes it necessary to attend to more than one variable in order to accurately classify system state.

Uncertainty was introduced by creating a borderline condition between pairs of state categories. A borderline condition exists when a range of values for one or more of the process variables simultaneously defines two system states. Thus, the borderline condition represents an area of uncertainty about the identity of an instance of system state.

Experimental Task

Subjects were trained as "operators" to classify the information presented on a display into one of four system state categories. The classification scheme required operators to integrate information from each of the four process variables. Accurate classification required that the operators learn the decision rules defining state category membership. To insure that the operator experienced the full range of exemplars from a state category, the frequency of instances selected from a category remained constant. Training consisted of classifying 64 instances of system data from each system state category for a total of 256 classification trials. Only operators who attained 90 percent correct classification accuracy were allowed to continue in the experiment.

The 20 Operators who reached criterion were then asked to rate the similarity of a subset of instances of system data. Four representative instances of system data from each state category were selected for inclusion in the pairwise similarity rating sessions. The four instances from each category (represented by a letter) are presented in Table 2. The instances were chosen to represent the full range of possibilities from that category. All possible pairs of these 16 instances were constructed (excluding identical pair combinations) and randomly ordered for presentation to the operator. There were a total of 120 pairwise combinations.

INSERT TABLE 2

ABOUT HERE

System State Representations

Two types of display representations were used to present system data: a separable display; and an integral display. The Digital display was a separable representation presenting the values of the four process variables as digits in a horizontal line. The Configural display was an integral representation of the same data, presenting the values of the process variables on a two-way axis. Examples of the two types of displays are presented in Figures 1a and 1b.

INSERT FIGURES 1a AND 1b

ABOUT HERE

Procedure

The experiment involved three sessions: 1) the Classification Training session; 2) the Post-Training similarity ratings session; and 3) the Extended Practice classification session. Each operator participated in the experiment for approximately 2 hours.

The Classification Training Session. Subjects were first briefed concerning the purpose of the experiment and were asked to sign an experimental consent form. Subjects were then randomly assigned to either the Configural or Digital display group. During training, the "operators" classified instances of system data. Presentation of each instance constituted a training trial. Each training trial followed the same pattern: presentation of an instance of system data; the operator's response; and feedback on the accuracy of the response. The cycle was repeated until each operator had viewed all 256 trials in their stimulus set.

Post-Training Similarity Ratings. After reaching criterion, pairwise similarity ratings were solicited for the 120 pairwise combinations presented in a random order. The pairwise similarity ratings were obtained by rating each of the 120 pairs of instances on a scale from 0 (very similar) to 10 (very dissimilar). The pairs of instances were presented to the operator in the from of the representation corresponding to his or her display type. Operators were instructed to use whatever criteria seemed appropriate.

Extended Practice. After a short break following the similarity rating session, operators performed a second classification session with the same stimulus set as used in the Training session. The stimuli were presented in a different random order and the operators received no feedback about their classification performance. The data from this session provided the performance measures for comparison to the results of the MDS analysis.

Data Measurement and Analysis

Two types of data were collected to investigate the operator's internal model and classification performance. Each type of data analysis is described in the following sections.

Multidimensional Scaling Data. The post-training similarity rating data for each operator was used as input for the MDS analysis using the KYST-2 scaling program. Spatial configurations were generated by KYST-2 to determine the dimensions used by each operator. Solutions were generated for one to five dimensions. The configurations which produced the best value of Stress were visually examined to determine the extent of clustering and interpret the dimensions. The result of that analysis was a list of the dimensions used by the group and by each operator.

Structural ratios and direct measures of distance were derived from the coordinate distances obtained with the MDS analysis. The structural ratio (as discussed in the Introduction) was used as a measure of the degree of categorical clustering of the instances of system data. Structural ratios for each of the dimensions identified in the visual inspection of the spatial configurations were calculated for each operator.

Since structural ratios are not sensitive to vector distances between instances, measures of directed distances were also calculated. The vector distance of interest to this research was Distance from Borderline. The starting reference point for the calculation of the measure were the four instances from each category located in the most proximate position to the borderline (instances D, H, L, and P). Within each category, the distance between that category's reference point and all other instances in the category were calculated. This produced three step sizes of distances from the borderline (e.g., the distance between pairs of instances D and C, D and B, and D and A in State Category 1). These distances were calculated for each State Category for each operator.

Performance Measures. The performance measures of interest in this study were accuracy and response time. Accuracy (percent correct classifications) for each operator was averaged over eight blocks of 32 trials in both the Training and Extended Practice sessions. The data from the Training session were used to assure that the operators had reached the 90% criterion performance level in the last 100 trials of that session. Only data collected from operators who reached the criterion were used in the analysis.

Response times were measured as the interval between the onset of an instance of system data and the operator's classification response. Response times were measured in milliseconds and only correct response times were used in the analysis; response times were averaged across eight blocks of 32 trials for each operator. The performance data were submitted to multifactor repeated measures ANOVA.

RESULTS

Interpretation of Group Dimensions

A summary of the Group and Individual MDS analyses is presented in Table 3. The Group results showed that, after training, the best fit for the Configural display group was in three dimensions and the best fit for the Digital display group was in four dimensions. The spatial configurations producing the best fit for the Group MDS analysis were visually inspected and labels assigned to the dimensions. The results of the analysis indicated that the dimensions used by the Configural display operators and the Digital display operators were different. The dimensions underlined in Table 4 are the interpreted dimensions for each display group.

INSERT TABLES 3 AND 4

ABOUT HERE

The dominant dimensions for the Configural display group were Orientation and Distance from Borderline. A representative, two dimensional graph of a Configural operator's spatial configuration is presented in Figure 2a. The figure clearly shows Orientation as the polarization between the two pairs of state categories (1 and 3, and 2 and 4) along the x-axis. Distance from Borderline (i.e., instances occupying the same relative position from the borderline) is represented by the y-axis; instances occupying the same relative position from the borderline (e.g., instances A, E, I and M) occupy the same relative position along that dimension in the spatial configuration. A third dimension (not shown) was also related to the shape of the stimulus; instances which were primarily "thin" (one side of the figure always significantly longer than the opposite side) were clustered at one end of the dimensions with "blocky" instances clustered at the other end of the dimension.

INSERT FIGURES 2a AND 2b

ABOUT HERE

The dimensions used by the Digital display group were distinctly different. Figure 2b presents a two-dimensional plot of two of the four dimensions revealed in the Group analysis. In fact, the figure shows three of the four dimensions: mid-range values to extreme values from left to right along the x-axis; values of process variables Q and M defining state categories 1 and 3 at the top of the y-axis, and values of B and H defining state categories 2 and 4 at the bottom of the y-axis; and four distinct areas in the geometric space corresponding to category membership (shown by the dotted lines in the figure). The fourth dimension (not shown in Figure 2b) corresponds to the values of Q and M; values of process variables Q and M defining state categories 3 and 4 are clustered at one end of the dimension while values of Q and M defining state categories 1 and 2 are clustered at the opposite end of the dimension.

Interpretation of Individual Operator Dimensions

Although the Group MDS analyses revealed a number of dominant dimensions, the Stress values indicated that the best fit to these data was not obtained. The Individual data in Table 3 shows that scaling each operator's data produced excellent Stress values, and a best fit in more dimensions. This result indicates that the Group MDS analysis did not completely reveal all the dimensions being used by operators in this task.

The spatial configurations for each operator were visually inspected and labels assigned to each dimension. All the dimensions identified across operators are presented in Table 4. Of the six dimensions identified for each of the two groups, Category Membership and Distance from Borderline best reflect the underlying statistical properties of the task. The remaining four dimensions represent attributes which are unique to the type of display. Thus, while operators in the Configural group were clustering instances of system state on the basis of "shape", "orientation", etc., Digital operators were using the dimensions of "mid/extreme values" and relative high and low values of process variables B and H.

It is interesting to note that clustering in the Configural group is similar to clustering in the Digital group despite the fact that the attributes are different. For instance, the Orientation dimension in the Configural group is the counterpart to the Digital dimension that separates instances based on relative high and low values of process variables B and H. Thus, while the attributes defining corresponding dimensions differ according to the unique properties of the displays, clustering of instances along those dimensions coincide. Similar correspondence was found for the other dimensions.

Notice that not all of the dimensions are continuous. In a number of cases, clustering occurs on the basis of discrete or nominal values along a particular dimension. This illustrates that the decision criteria used by the operator to judge similarity may not be based strictly on a continuous scale, but based on the presence or absence of a specific value or set of values for a given attribute.

Structural Ratios as a Measure of Clustering

To determine if a particular dimension dominated an operator's conceptual space, structural ratios were calculated. If one assumes that the dimension producing the minimum structural ratio is the dominant dimension for that operator, then the results clearly illustrate the difference in emphasis between operators. For example, Category Membership is the dominant dimension for three of the Configural operators and for two of the Digital operators. Here, however, the similarity ends. The Orientation and Exact Shape dimensions are dominant for the remainder of the Configural group, whereas the rest of the Digital group found the Distance from Borderline and the Mid/Extreme dimension to be the most important. These results suggest that operators can perceive system data in significantly different ways, especially among the Digital display operators in this experiment.

Direct Measures of Distance

Judgments of similarity based on dimensions other than category membership (such as Distance from the Borderline) will not be reflected in the structural ratio. In such situations, the position of an instance in geometric space can be captured only by a proximity measure which characterizes the relative distance between instances along the dimension of interest. For example, if Distance from the Borderline is a dominant dimension for the operators, then instances occupying positions in the category that are distant from the overlap region between state categories should be proportionally distant in the geometric space.

To test whether Distance from the Borderline became a dominant dimension for operators after training, the coordinate values from the MDS analysis were used to calculate directed distance measures. Mean distances for the three step sizes from the borderline for the Configural display operators and the Digital Display operators are presented in Figure 3.

INSERT FIGURE 3

ABOUT HERE

For the Configural display operators, distance increased as the position of instance within a state category away from the borderline condition. A similar relation between distance and position from the borderline is not evident in the Digital display group. An ANOVA of these data showed that

Distance from the Borderline, $F(2,36) = 11.07$, $p = .0002$, and the Distance from Borderline by Display type interaction, $F(2,36) = 26.48$, $p < .0001$, were highly significant. An analysis of the simple main effect of Distance from Borderline for the Digital display group found the differences between step sizes not to be significant, $F(2,18) = 2.39$, $p = .1203$. These results confirm the previous analysis of dimensions; in general, Distance from Borderline was a dominant dimension for the Configural display operators, but not for the Digital display operators.

Evaluating the slope of the function relating the scaled distances to Distance from Borderline position reveals the dominance of that dimension for the Configural display operators. A multiple regression analysis of the data revealed a correlation of $r = 0.56$ between borderline distance measures and scaled distances. The positive trend is evident for all Configural operators but one. Since that operator's data points were more than 2.5 standard deviations from the mean, his data were excluded and the data reanalyzed. The correlation for the corrected analysis increased to $r = .84$, indicating a strong positive relation between Distance form Borderline and scaled distance measures for virtually all of the operators in the Configural display group.

As might be expected, the correlation between scaled distances and Distance from Borderline was very low ($r = .16$) for the Digital display group. Inspection of the slopes of the regression equation for this group provide a number of interesting insights into the effect of other dimensions on the scaled distance measure in this analysis. For instance, the downward trend evident in the distance measures of some of the Digital operators may be the result of an interaction between dimensions. Since Distance from Borderline is not dominant for the Digital group, the interaction with the Mid/Extreme dimensions would cause the distance between the closest and farthest borderline positions to be less than the distances between the closest and middle two borderline positions.

Relating Dominant Dimensions to Performance

The importance of Distance from Borderline to the two groups of opera-s was mirrored by their performance. Figure 4 shows mean response times for the Digital and Configural display operators as a function of Distance from the Borderline during the Extended Practice session. Analysis of these

---

INSERT FIGURE 4

ABOUT HERE

---

data showed that response times for the Configural group significantly decreased as Distance from the Borderline increased; whereas the response times for the Digital display group were not significantly effected by Distance from Borderline. The ANOVA confirmed these results. Distance from the Borderline was found to be highly significant, $F(6,102) = 15.98$, $p < .0001$, as well as the Distance from Borderline by Display type interaction, $F(6,102) = 2.59$, $p = .0223$. Analyses of the simple main effects of display type found a significant main effect for Distance from the Borderline for the Configural display, but not for the Digital display.

When Figure 4 is compared to Figure 3, a clear trend emerges for the Configural display; as distance from the borderline increases, response times decrease. Thus, response times are inversely related to Distance from the Borderline for the Configural group operators. Such a relation is not evident for the Digital group, indicating that this dimension was relatively unimportant in determining overall operator performance.

To find a more direct relationship between classification performance and the dominance of the Distance from the Borderline dimension, the difference in response times for those instances used in the MDS analysis were calculated and compared to the slope of the distance function for each operator. For example, the difference in response times between instances a and b, a and c, and a and d represent the data for State Category 1. Differences for the three points were averaged across state categories, providing three response time data points for each operator. Directed distances were calculated in the same way. To assess the magnitude of the change of the slope, confidence limits were established for each operator's data using the Tukey(a) procedure. If Distance from the Borderline dimensions reflects the operator's internal conceptual model, then the slope of the difference in response times should correspond in some way to the slope of the distance measures.

For the Configural group, those operators whose response times did not differ significantly across borderline conditions had an average slope of 0.156; operators whose response times differed significantly had an average slope of 0.394. In the Digital group, operators producing a significant difference in response times had an average slope of -.028, with those operators exhibiting no significant difference in response times had an average slope of -.081.

CONCLUSIONS

The primary goals of this exploratory research were to show that multidimensional scaling is an effective methodology for: 1) revealing the underlying decision criteria used by operators to classify system state; 2) determining the impact of display representation on the development of the operator's internal model during learning; and 3) relating the operator's internal model to performance. In general, the results reported in this paper show that MDS is an effective and useful tool for evaluating the structure and composition of an operator's mental model, and an effective method for assessing the impact of display representations on those internal models.

Composition of the Operator's Model

An important result of this study is that the information used by the two display groups to identify system state was qualitatively different. The comparison of the spatial configurations derived from the MDS analysis after training showed that the dominant dimensions for the Configural display group were different than the dominant dimensions for the Digital display group. For instance, Distance from the Borderline become a dominant dimension for the Configural display operators, but not for the Digital display operators.

The Digital display operators, on the other hand, placed a greater emphasis on the actual values of the process variables. For instance, the pairing of the state categories based on the values of process variables B and H emerged as one of the dominant dimensions during training. In addition, the Digital operators also grouped instances based on the values of process variables Q and M.

Despite differences in the composition of the operators' internal model, the structure of that model appeared to be relatively stable across operators. The evidence for such a conclusion can be found in the correspondence in the dimensions used by the two groups. For every cluster of instances found in the Configural group, there is a corresponding cluster of instances in the Digital group. Thus, both groups learned the same underlying correlational structure of the task, but mapped different attributes to state categories. This is not altogether surprising since the attributes of the displays are driven by the same underlying statistical properties; i.e., the relationship between attributes is the same in both displays since changes in the values of the process variables occur in a correlated fashion. One can conclude, then, that the structure of the operators' mental model in this task is the same, but composed of different information about attributes defining state category membership.

Therefore, we can conclude that the form of the display representation has a major impact on what it is that operators learn about state category membership. The choice of a display for representing system data becomes complicated by the fact that the display will significantly influence the composition of the operator's internal conceptual model of the task. In addition the results provide further evidence that the statistical properties of a decision task and the physical representation of system data are two important, but fundamentally different, issues in decision making. Previous research has tended to side-step that issue, creating experimental situations where the results may be confounded by the interaction between the underlying correlational structure of the task, and the physical representation of the data.

Relating Models to Performance

The fact that MDS can identify the structure and composition of the operator's internal model is an important result. The ultimate utility of the method is, however, dependent upon the relation between the operator's model and performance, and the ability of the results of an MDS analysis to predict operator performance. It was quite gratifying to see in this study that MDS was sensitive to the dimension which had the greatest impact on the performance of the Configural group. One can conclude, then, that Distance from the Borderline was the dominant dimension for the Configural display operators, and constitutes the primary underlying decision criterion used by those operators to process system data and classify system state in this task.

Although the relation between the Digital display operators' internal model and their performance is not altogether obvious, one can speculate about possible decision strategies used by that group of operators based on the dimensions revealed by the MDS analysis. For instance, the Digital display operators could have been using a two-step classification procedure: first, reduce the decision problem from four alter 'ves to two alternatives by attending to process variables B and H. Once the most likely pair of state categories is identified (1 or 3; 2 or 4), then attend to process variable Q or M. In situations of relatively low uncertainty, that strategy would equivalent response times. Processing times increase only when circumstances dictate sampling another process variable (for instance, under conditions of high uncertainty). Therefore, response times become strictly a function of the number of process variables the operator has to sample to reach a classification decision. The importance of distance from the borderline condition would become evident only in those situations where fewer variables had to be sampled. Thus, MDS may reveal the significant dimensions being used by an operator, but not be sensitive to the subtle influences on performance that will occur with certain types of classification strategies.

Individual Differences

An important finding of this research is that the differences between individual operators was quite significant. In fact, these results suggest that focusing on an aggregate analysis and ignoring the unique way in which individual operators internally represent decision criteria can be dangerous.

Fortunately, MDS provides the data necessary for assessing the disparity between the Group analysis and individual operator dimensional configurations. For instance, Distance from the Borderline was a very dominant overall dimension for the Configural operators, but the importance of that dimension varied between operators. Without ascertaining the importance of a dimension for an operator, assumptions about the significance of a dimension may not be valid. This is especially important to designers of decision aiding systems; without a method for assessing the salience of decision criteria for all operators, there is no way to assure that the model being used by the decision aid is congruent with the operators internal conceptual model of the system. This research showed that MDS can be used to test those assumptions.

In addition, MDS may aid in the design of adaptive decision aiding systems. Customizing an interface to suit individual operators presupposes that the decision aid know something about each operator's mental model. MDS can provide information about the composition and structure of those models so that the decision aid will understand how a particular operator weights the importance of certain decision criteria.

In summary, MDS has been shown in this research to be a potentially useful tool for assessing the internal models used by operators to represent state category information, and relate that model to performance. In addition, the methodology provides insight into the way in which a particular type of display can affect an operator's conceptual model.

ACKNOWLEDGEMENTS

The authors would like to thank Warren S. Torgerson for his most useful advise during the early formulation of this research. We are also indebted to the Department of Industrial Engineering and Operations Research at the University of Massachusetts for providing the facilities necessary for conducting this research.

REFERENCES

[1] J. Rasmussen, Information processing and human-machine interaction. New York: North-Holland, 1986.

[2] J. Rasmussen, "The human as a system component", in Human Interaction with Computers, H.T. Smith and G.R.E. Green, Eds. New York: Academic Press, 1980.

[3] J. Rasmussen, "Models of mental strategies in process plant diagnosis", in Human Detection and Diagnosis of System Failures, J. Rasmussen and W.B. Rouse, Eds. New York: Plenum Press, pg. 241-258, 1981.

[4] J. de Kleer and J.S. Brown, "Assumptions and ambiguities in mechanistic mental models", in Mental Models, D. Gentner and A.L. Stevens, Eds. Hillsdale, NJ: Lawrence Erlbaum Assoc., 1983.

[5] W.R. Garner, The Processing of Information and Structure. New York: John Wiley & Sons, 1974.

[6] E.E. Smith, and D.L. Medin, Categories and Concepts. Cambridge, MA: Harvard University Press, 1981.

[7] J.H. Holland, K.J. Holyoak, R.E. Nisbett, and P.R. Thagard, Induction: Processes of Inference, Learning, and Discovery. Cambridge, MA: The MIT Press, 1986.

[8] W.S. Torgerson, Theory and methods of scaling. New York: John Wiley & Sons, 1958.

[9] R.N. Shepard, "The analysis of proximities: multidimensional scaling with an unknown distance function I", Psychometrika, 27, 125-140, 1962.

[10] J.B. Kruskal, "Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis", Psychometrika, 29, 1-27, 1964.

[11] J.B. Kruskal, and M. Wish, Multidimensional Scaling. Beverley Hills, CA: Sage Publishing, 1977.

[12] S.K. Reed, "Pattern recognition and categorization", Cognitive Psychology, 3, 382-407, 1972.

[13] D.E. Rumelhart and A.A. Abrahamson, "A model for analogical reasoning. Cognitive Psychology, 5, 1-28, 1973.

[14] D. Homa, D. Rhoads, and D. Chambliss, "Evolution of conceptual structure", Journal of Experimental Psychology: Human Learning and Memory, 5(1), 11-23, 1979.

[15] B.G. Coury, "Multidimensional scaling as a method for assessing internal conceptual models of inspection tasks", Ergonomics, 30(6), 959-973, 1987.

[16] C.D. Wickens, "The Object Display: Principles and a Review of Experimental Findings", Cognitive Psychophysiology Laboratory Technical Report No. CPL 86-6, University of Illinois, 1986.

[17] E.J. Casey, "Visual display representation of multidimensional systems: the effect of information correlation and display integrality", in the Proceedings of the Human Factors Society 30th Annual Meeting, The Human Factors Society, 430-434, Oct. 1986.

[18] B.G. Coury, M.D. Boulette, M.C. Zubritzky and D.L Fisher, "Classification of multidimensional data: evaluating displays to aid identification of system state", in, Proceedings of the 1986 IEEE International Conference on Systems, Man and Cybernetics, IEEE SMC, 1503-1508, Vol. 2, Oct. 1986.

[19] M.D. Boulette, B.G. Coury and N.A. Bezar, "Classification of multidimensional data under time constraints: evaluating digital and configural display representations," in the Proceedings of the Human Factors Society 31st Annual Meeting, The Human Factors Society, 116-120, Oct. 1987.

[20] B.G. Coury and M.D. Boulette, "The impact of uncertainty on classification of multidimensional data: a comparison of integral and separable displays", Human Performance Laboratory Technical Report No. BGC1-1987, University of Massachusetts/Amherst, 1987.

Table 1

Ranges of Values for Process Variables

System State	Q	M	B	H
	Process Variable			
1	25-51	49-75	0-26	74-100
2	25-51	49-75	74-100	0-26
3	49-75	25-51	0-26	74-100
4	49-75	25-52	74-100	0-26

Table 2

Category Position of Instances

Selected for Pairwise Similarity Ratings

System State	3	2	1	Border
	Distance From Borderline			
1	A	B	C	D
2	E	F	G	H
3	I	J	K	L
4	M	N	O	P

Table 3

Summary of Multidimensional Scaling

Analysis of Similarity Ratings

	Pre-Training		Post_Training	
	#dim	Stress	#dim	Stress
Configural				
Group:	3	.180	3	.185
Individual:	5	.037	4.2	.030
Digital				
Group:	3	.189	4	.165
Individual:	4.6	.026	4.5	.028

Table 4

Dimensions Identified from the Spatial Configurations

for the Configural and Digital Display Groups

Configural	Digital
category membership	category membership
borderline position	borderline position
approximate shape	middle/extreme values(4 groups)
orientation	categories 1 and 3; 2 and 4
line vector between Q,M	same Q and M
exact shape	mid/extreme values(8 groups)

Table 5

Structural Ratios for Each Operator's Dimensions

Identified in the Spatial Configurations

Configural Display Group

Sub	Cat mem	Borderline	Orientation	Q and M	Approx shp	Exact shp
1	0.641	-----	0.559*	1.391	-----	1.283
2	0.562	-----	0.550*	1.311	-----	1.256
3	0.781	0.765	-----	0.990	-----	0.629*
4	0.657*	1.034	0.670	-----	-----	-----
5	0.841	-----	0.963	-----	-----	0.608*
6	0.751*	0.954	0.775	-----	-----	0.857
7	0.735*	-----	0.943	-----	-----	-----
8	0.942	0.732	-----	-----	-----	0.325*
9	0.887	0.772	-----	-----	1.056	0.689*
10	0.772	0.992	0.908	-----	-----	0.676*

Digital Display Group

Sub	Cat mem	Borderline	SS1&3;2&4	Q and M	Mid/Ext:4	Mid/Ext:2
11	-----	0.368*	-----	0.384	-----	0.528
12	0.843	0.948	0.858	-----	0.829*	-----
13	-----	0.992	0.891*	-----	-----	-----
14	1.320	-----	-----	1.112	0.487*	-----
15	0.669	0.028*	-----	0.037	-----	0.041
16	0.875	0.876*	-----	-----	-----	-----
17	0.735	0.344*	0.694	0.388	-----	-----
18	1.205	0.752	-----	0.720	0.536*	-----
19	0.302*	0.501	0.724	0.690	-----	-----
20	0.869*	-----	0.885	-----	-----	1.088

* indicates the dominant dimension for each operator

Figure Captions

Figures 1a and 2b. Reproduction of the Digital display (Figure 1a) and the Configural display (Figure 1b) with each showing the same instance of System State Category 3.

Figures 2a and 2b. Representative 2-dimensional spatial configurations for a Configural display group operator (Figure 2a) and for a Digital display group operator (Figure 2b).

Figure 3. The relation between an instances' position in a state category relative to the borderline conditions and scaled distances from the borderline for the Configural and Digital display.

Figure 4. Response times to the Configural and Digital displays as a function of Distance from the Borderline condition.

Figure 1a

Q	M	B	H
66	34	09	91

Figure 1b

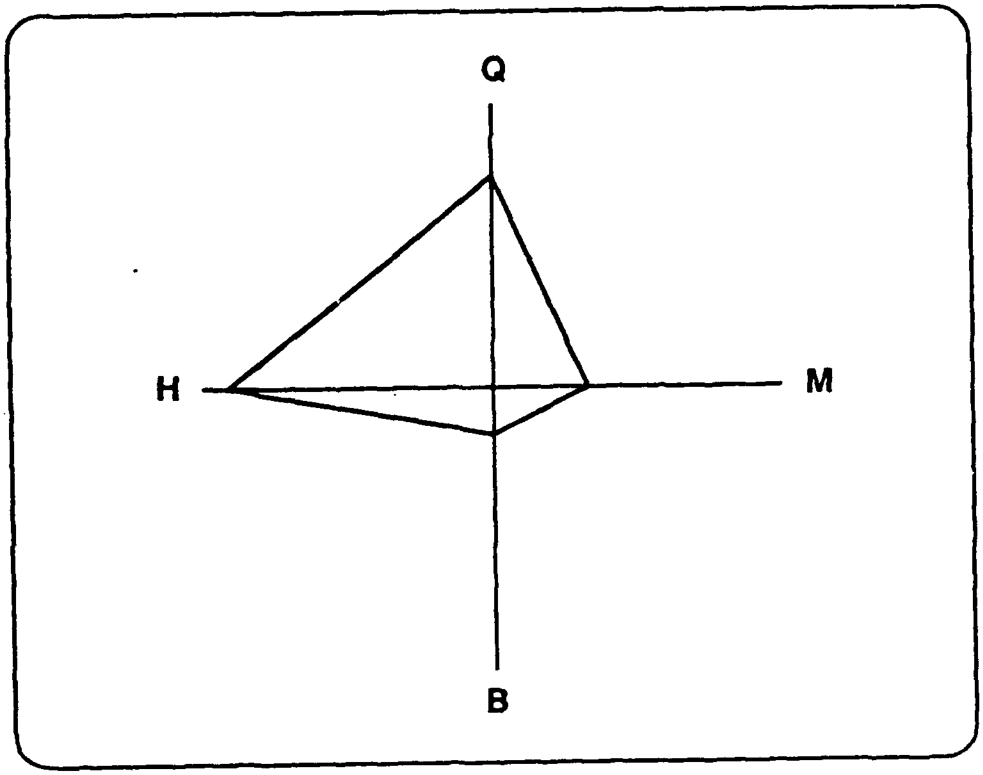


Figure 2a.

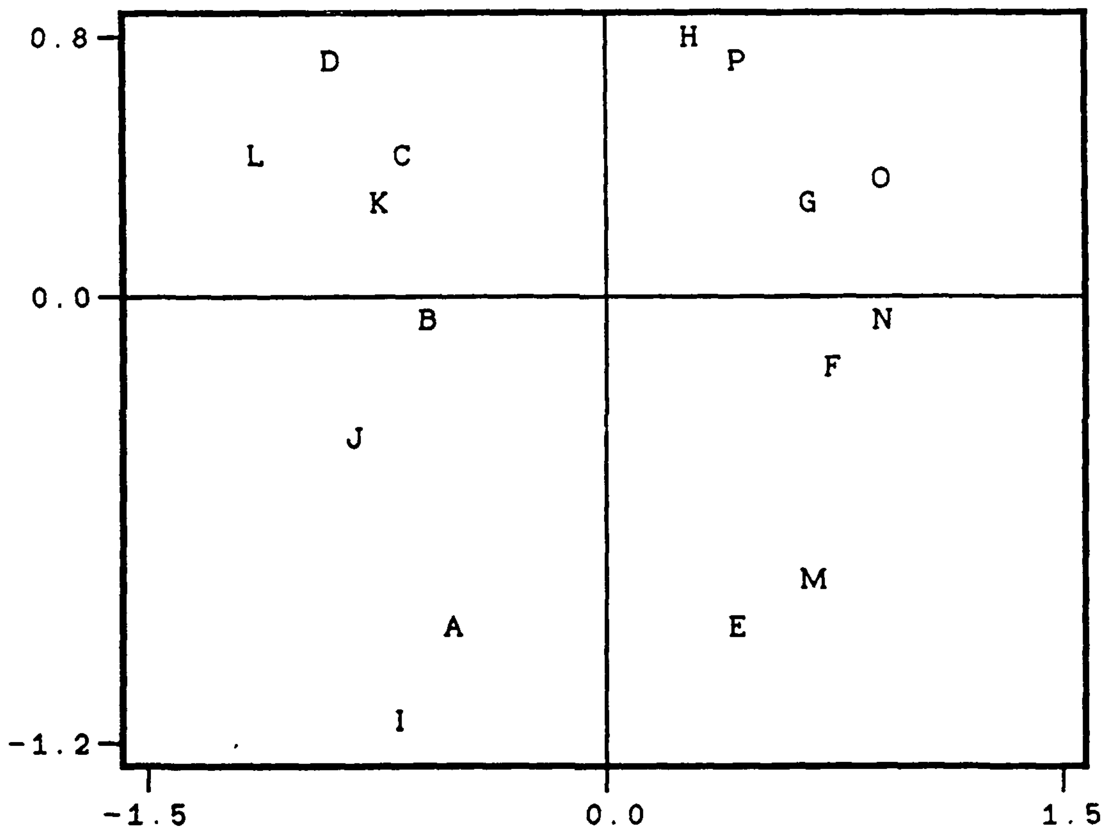


Figure 2b

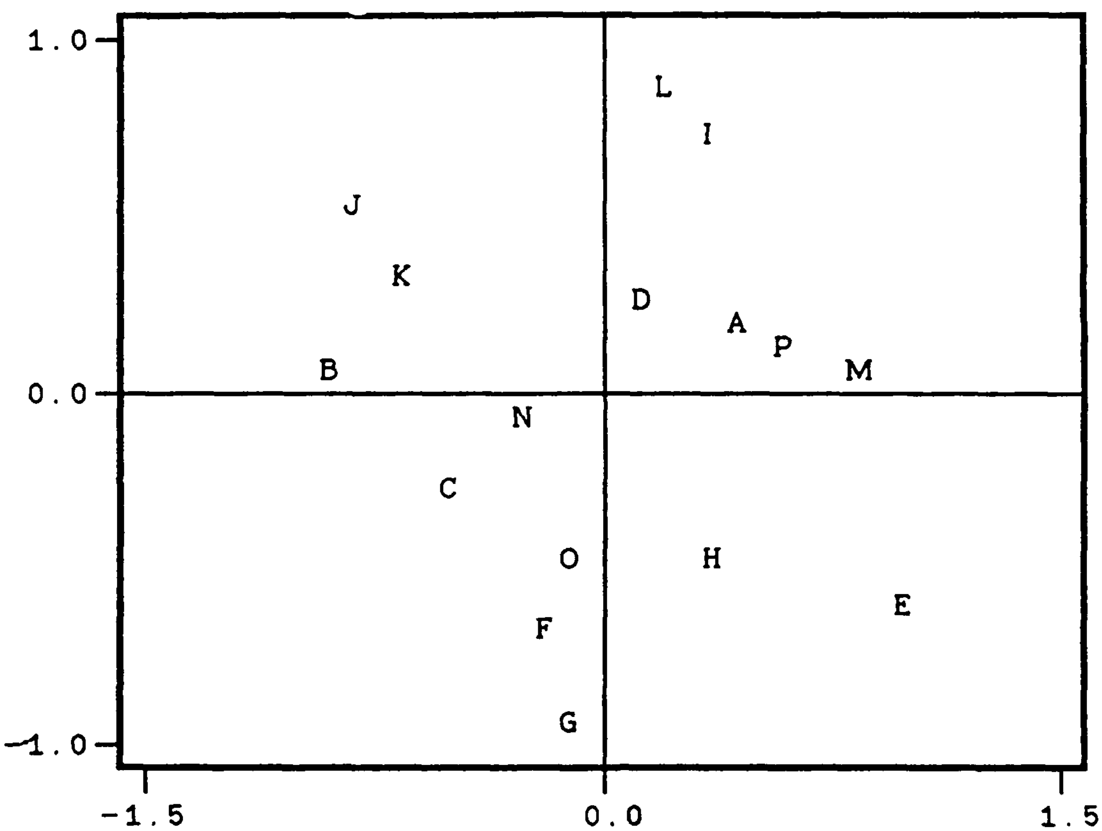


Figure 3

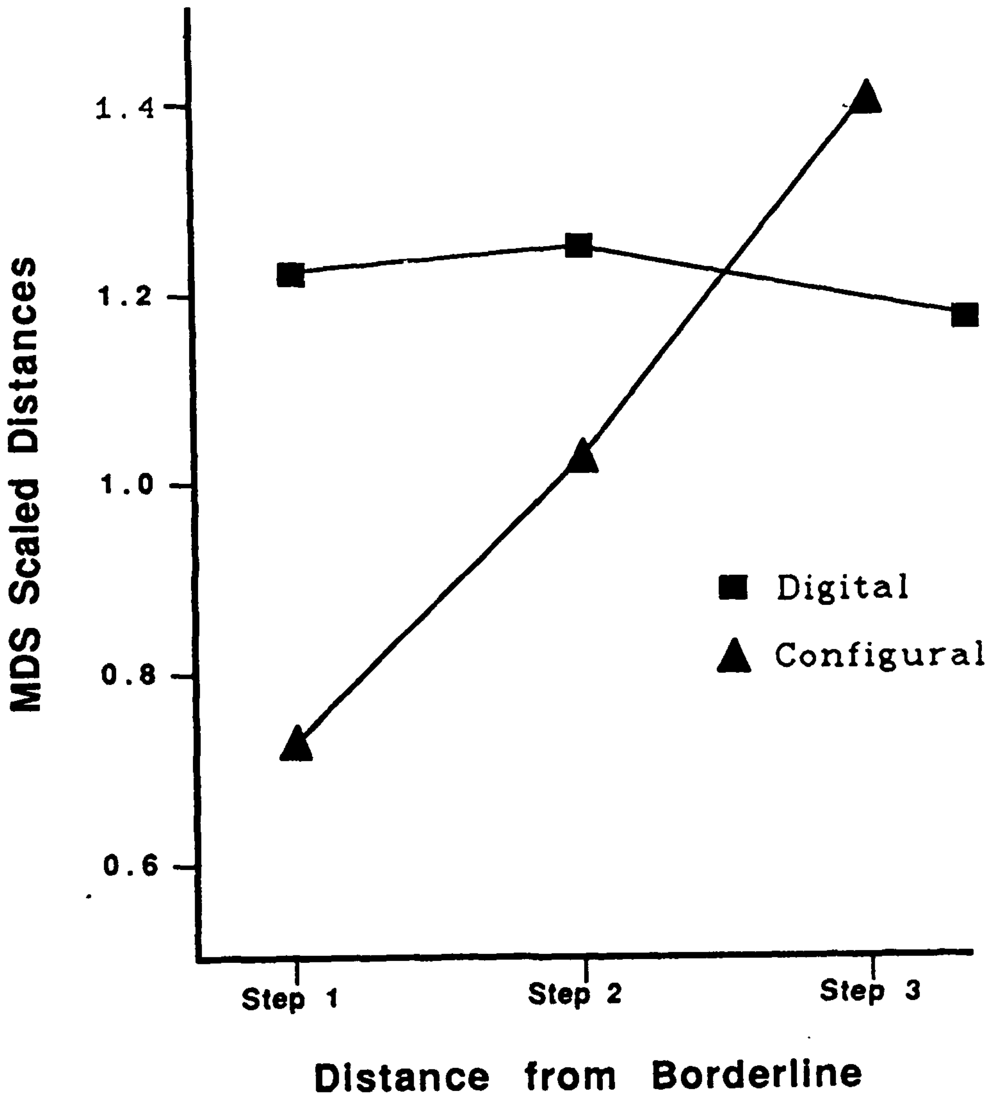
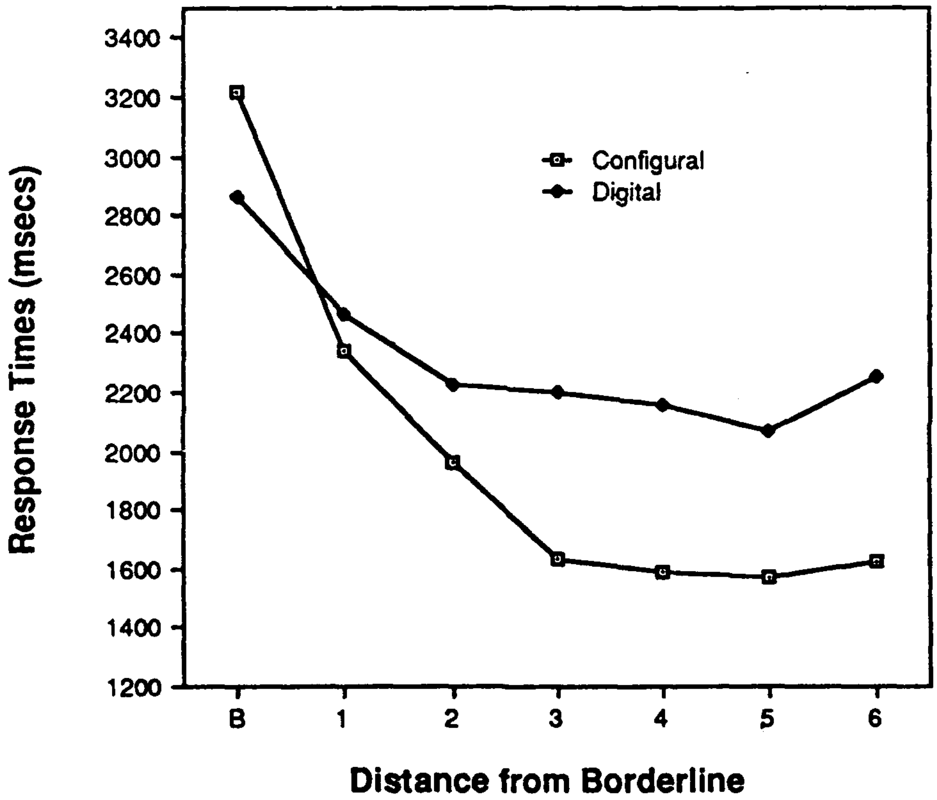

MDS Scaled Distances
1.4
1.2
1.0
0.8
0.6
Digital
Configural
Step 1
Step 2
Step 3
Distance from Borderline

Figure 4



Configural
Digital
Response Times (msecs)
3400
3200
3000
2800
2600
2400
2200
2000
1800
1600
1400
1200
B
1
2
3
4
5
6
Distance from Borderline